The Case for Using Semantic Nets as a Convergence Format for Symbolic Information Fusion

Jean Rohmer
Thales Communications
160, Boulevard de Valmy
BP 82
92704 Colombes Cedex
FRANCE

jean.rohmer@fr.thalesgroup.com

ABSTRACT

We explore how various information formats can be merged into an unique semantic space, using the Semantic Nets formalism. We show that this formalism can then be transformed and reworked to perform classical data analysis computations, which will help in the fusion and discovery process. We advocate for using semantic nets to get sense from heterogeneous informations, in particular texts, as a step towards what could be called "Litteratus Calculus".

Keywords: *Symbolic Information Fusion, Semantic Nets, Data Analysis, Text Mining, Text Understanding, Automatic Classification.*

1.0 INTRODUCING SEMANTIC NETS AS A WAY TO COLLECT HETEROGENEOUS INFORMATIONS INTO AN UNIQUE FORMAT

A key issue in information fusion is to deal with very different natures of data : numerical data, usually in the form of simple tables, more complex structured data like relational databases, semi-structured messages, totally unstructured texts. Moreover, efforts in information standardization recently introduced new formats like XML and its many derivates, TOPIC MAPS, UML models, ontologies models.

We have to deal with an impressive continuum of representations, from fully numeric and structured to totally textual and unstructured.

Solving this situation of heterogeneity is a prerequisite to information fusion processes and algorithms.

In general, Intelligence Information System designers are facing a difficult choice :

- either adopt a structured approach, e.g . choose to unify their data in a large relational database
- or, in the opposite direction, keep all the information under the form of documents

In practice, each model excludes the other one : in the first case, information will be accessed through structured query languages, along with programming of specific applications to interface the user with data.

In the second case, only text search engines are available to retrieve documents containing the desired piece of information. Attempts to put together relational and textual paradigms usually lead to costly and uncomfortable designs.

Paper presented at the RTO IST Symposium on "Military Data and Information Fusion", held in Prague, Czech Republic, 20-22 October 2003, and published in RTO-MP-IST-040.

Report Documentation Page

Form Approved
OMB No. 0704-0188

Public reporting burden for the collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington VA 22202-4302. Respondents should be aware that notwithstanding any other provision of law, no person shall be subject to a penalty for failing to comply with a collection of information if it does not display a currently valid OMB control number.

1. REPORT DATE **00 MAR 2004**	2. REPORT TYPE **N/A**	3. DATES COVERED **-**

4. TITLE AND SUBTITLE **The Case for Using Semantic Nets as a Convergence Format for Symbolic Information Fusion**	5a. CONTRACT NUMBER
	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Thales Communications 160, Boulevard de Valmy BP 82 92704 Colombes Cedex FRANCE**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release, distribution unlimited

13. SUPPLEMENTARY NOTES
See also ADM001673, RTO-MP-IST-040, Military Data and Information Fusion (La fusion des informations et de données militaires)., The original document contains color images.

14. ABSTRACT

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT **unclassified**	b. ABSTRACT **unclassified**	c. THIS PAGE **unclassified**	**UU**	**34**	

Standard Form 298 (Rev. 8-98)
Prescribed by ANSI Std Z39-18

However, in an human brain, this distinction between structured and unstructured data simply does not exist : we, as humans, are able to merge informations coming from a newspaper, an Excel file, a database, an oral conversation …

How can computers mimic our extraordinary capability to make information fusion in our brain ? The solution is to represent information in the machines in a way not too far from the way it may be represented in our heads. This subject has been studied for years in the field of « Artificial Intelligence », and, as early as in the 50's, came the concept of *Semantic Nets* to meet this challenge.

Although in the 80's Artificial Intelligence applications were disappointing to a point that this field of computer science was nearly abandoned in the 90's, we have made recently the proof that Semantic Nets, the representation side of AI –as opposed to its automatic reasoning side- was both an extremely efficient and human friendly way of representing complex informations.

We started developing and using in the early 90's a tool dedicated to the management of Semantic Nets, IDELIANCE. Today we can confirm that Semantic Nets is a practical and efficient way of handling heterogeneous information sources.

Ideliance was originally designed as a personal knowledge management tool. The initial idea is to offer an information representation model which bridges the gap between structured data (like tables in spreadsheets and relational databases) and unstructured data (found in documents written in natural language). Semantic Nets appear to be a nice compromise between data and texts.

They can be viewed as a collection of simple sentences « Subject / Verb / Object » :

Peter / works for / Mary
Mary / lives in / Berlin

A key property in Ideliance is that each sentence is represented in both directions :

Mary / employs / Peter
Berlin / is the place where lives / Mary

Subjects can also be long sentences identifying a complex but precise concept :

The 23rd March 2003 ACM Meeting in Berlin about the ALPHA project

With Semantic Nets, a solution is to forge as many S / V / C sentences as necessary :

Berlin / is the place of / The 23rd March 2003 ACM Meeting in Berlin about the ALPHA project

ACM / is the organiser of / The 23rd March 2003 ACM Meeting in Berlin about the ALPHA project

ALPHA Project / is addressed at / The 23rd March 2003 ACM Meeting in Berlin about the ALPHA project

…

Other pieces of knowledge may be expressed this way :

German chapter of ACM / is located in / Berlin

German chapter of ACM / is member of / ACM

Note that *ACM* , *German Chapter of ACM* and *ALPHA Project* are themselves *Subjects.*

In an Ideliance collection, dizains of thousands of subjects can be found. By constrat, verbs will generally amount only to few to many dizains. They represent the vocabulary which describes the domain of the application.

Interestingly, it becomes easy to know « everything about Berlin » :

Berlin / is the place of / The 23rd March 2003 ACM Meeting in Berlin about the ALPHA project

Berlin / is the place where lives / Mary

Berlin / is the place of / German Chapter of ACM

(Each of these sentences may come from a different source and / or format : text, database, message, …).

When clicking on a given subject, a page is built with all the sentences starting with it. Navigation continues by clicking on one of the subjects at the end of these sentences.

Ideliance can be seen as a general purpose tool for managing such sets of sentences, applicable in many real life situations by non specialists in computer programming. It exists under the form of a personal tool on Windows, and of an HTTP server. Ideliance is in operation for more than three years in various application contexts : military intelligence, knowledge management, competitive intelligence, experience sharing among teams.

Users can edit new sentences through graphical editors, either by reusing the existing vocabulary, or by creating new subjects or verbs. Statements can also be obtained by *automatic translation of structured data* (Excel, SQL, XML …) into sentences. Text mining tools outputs can also be translated into Ideliance sentences.

N.B. The old concept of semantic nets has recently been found in knowledge representation tools developped in the context of Internet. Formalisms for ontologies representation (the coming W3C OWL standard), and more general information representation (the W3C RDF standard notation, based upon XML) are proposed by the Internet community under the general *Semantic Web* umbrella, with the vision that, in the future, information on the Web should be written in such a formalism rather than in textual pages. A tool like Ideliance, dedicated to Semantic Nets management, can be seen as « Semantic Web *avant la lettre* », and also as a practical tool to run dedicated « Semantic Intranets » without waiting for the hypothetical rise of the Global Semantic Net.

2.0 FROM FORMAT FUSION TO INTELLIGENCE FUSION

We can now address the core topic of this paper : once heterogeneous data (from documents, databases, messages, spreadsheets, …) have been gathered in a Semantic Net format, how to process this net to achieve *fusion* ?

Vocabulary remark : for most information technology people, converting various formats and databases into an unique format, is itself called « fusion » : we started from five databases, we end up with just one. We call this kind of fusion « **Format Fusion** ».

For Intelligence people, information fusion (we will call it **Intelligence Fusion**) is a totally different concept : we receive 20 documents or messages describing riots in a town at a given date, we ask ourself the following questions :

- are all these messages about the same unique riot, or do they report about several ones ?
- if there is an unique riot, how to identify and characterise the elements which describe it : number of participants, mode of action, consequences on the population, damages caused … Again, several messages will deal with damages : are they referring to the same damage, or to several ones ?

- if there are several riots, how to distinguish them –i.e. how to refer to them, to name them ? Some papers may be about only one of theses riots, other papers will deal with several ones. And the question of delineating each of the characteristics of each riot raises as well.

Stated like this, **Intelligence Fusion** appears to be much more complex than **Format Fusion**. However, the better the Format Fusion process will be conducted, the better the Intelligence Fusion process will start. In one hand, Format Fusion is a prequisite to deploy automated, computerized procedures for Intelligence Fusion. In another hand, Format Fusion can already help « manual », « human » Intelligence Fusion, simply by offering an unified, seamless way to navigate, browse through the whole set of informations, collected in a unique semantic net.

It is clear that Intelligence Fusion, in general, is extremely complex and difficult.

It can be tempting to use « brute force » to solve it : start from messages texts, do some form of terminology extraction, (text mining), then run statistical tools or « business intelligence » tools. We then face many problems, among which :

- how to differentiate by statistics 20 messages about 2 different riots, or 15 messages about 10 different riots
- if one message mentions 25 casualties, another one 45 casualties, are they figures about two diffrent riots, or about the same one ? In this case, what to do with these two figures : take the minimum, the maximum, take their sum, their average ?
- when a riot is mentioned in a message, does it concern a new incoming event, or is it a reference to a past event ?
- how to take into account the bias followed by the authors of the messages ?

These huge problems may advocate for the need of a fine grain analysis of the natural language used in the documents, including tenses, conditional modes, *nuances,* ...Unfortunately, current state of the art in natural language understanding and interpretation is far behind what is needed here.

Our experience suggest the following steps:

a) Starting from documents, use text minig tools and terminology extraction tools to *prepare the documents for semantic modelling*

b) Translate – mainly « by hand », with the help of tools like Ideliance – the preprocessed documents into an unique semantic net. This realises **Format Fusion**.

c) Perform **Intelligence Fusion** – both manually and automatically – on the unique resulting semantic net

N.B : Automatic translation of structured informations into Semantic Nets is not difficult, since, by definition, the semantics of structured data is known with precision. (Ideliance, for instance offers severals tools to automatically translate spreadsheets and relational databases into semantic nets). It is thus easy to inject structured information (e.g. about geography, weapons) into the semantic net.

3.0 SOME BASIC MECHANISMS FOR INTELLIGENCE FUSION IN A SEMANTIC NET

3.1 How to Formalise the Fusion Problem

We consider now that we start with a set of informations represented in a semantic net. We consider that the Intelligence Fusion process has not yet been processed.

That means for instance that, if we, at the beginning started with :

- message 1, mentioning an event with one person and one car
- message 2, mentioning an event with two persons and two cars

we have created the following subjects :

Event 1, Event 2, Person 1, Person 2, Person 3, Car 1, Car 2

Some sentences are also created, such as :

Person 1 / is mentioned in / Event 1

Car 1 / is mentioned in / Event 1

Person 2 / is mentioned in / Event 2

Person 3 / is mentioned in / Event 2

Car 2 / is mentioned in / Event 2

At this point, it is important to note that :

We do not know if Event 1 and Event 2 are the same or not, (idem for Car 1 and Car 2)

We know that Person 2 is different from Person 3, but each of them may be the same as Person 1

Formally, we can see each subject in our semantic network as a ***variable***, along with constraints (equations, inequations) about theses variables, and we can represent these constraints themselves as sentences in the network :

Person 2 / is different from / Person 3

Event 1 / may be equal to / Event 2

Car 1 / may be equal to / Car 2

Person 1 / may be equal to / Person 2

Person 1 / may be equal to / Person 3

The objective of what we call Intelligence Fusion is to reduce uncertainty, i-e :

- to conclude that some variables are the same
- to conclude that some variables are not the same

As for any system of equations, we need some ***constants*** to ground the system, and to bootstrap the solving process.

Event 1 / takes place / Avenue des Champs Elysées

Event 2 / takes place / Rue de Rivoli

This could lead to the conclusion that :

Event 1 / is different from / Event 2

But certainly not that *Car 1 / is different from / Car 2 !*

3.2 Similarities and Differences, Identities and Distances

We are here in the domain of symbols (a street name, a person name) rather than the domain of numbers (a speed, a pressure, a geometric position).

We have to deal with similarities and differences in a symbolic, discrete world, not in a numerical, continuous world.

Whereas in the latter case, the key point is the –continuous- notion of *distance*, in our case, the notion of ***identity*** prevails.

« Rue de la Paix » is not identical to « Place Charles de Gaulle »

« Rue de la Paix » is identical (only) to … « Rue de la Paix »

In other words, we advocate here, -due to the complexity of the problem- ***not to*** try to transform the symbolic, discrete world, into a continuous world (through fuzzy sets, bayesian networks, possibilities …). There are already enough progresses to do to adress Intelligence Fusion in a discrete symbolic world. (We conjecture that human judgments and decisions –the ultimate goal of Intelligence Fusion output—are more discrete than continuous : « I choose A against B », « I think that James is a nice guy ».

And discrete symbols are capable of describing details and *nuances* :

Given the sentences :

Event 1 / takes place / Rue de Rivoli

Event 2 / takes place / Place Charles de Gaulle

We can add that :

Rue de Rivoli / is located in / Paris 2ème

Avenue des Champs Elysées / is located in / Paris 8ème

Paris 8ème / is member of / Paris Luxury Districts

Paris 1er / is member of / Paris Luxury Districts

We see that ***Event 1 and Event 2 share a common point :***

They both happened in streets belonging to a district among the Paris Luxury Districts

More precisely, we say that two subjects SA and SB have a point in common if there exist sequances of sentences of the form :

(SA / **V0** / SA1) (SA1 / **V1** / SA2) (SA2 / **V2** / SA3) … (SAn **VN** **S**)

(SB / **V0** / SB1) (SB1 / **V1** / SB2) (SB2 / **V2** / SB3) … (SBn **VN** **S**)

(where all SA, SAi are different, and all SB, SBi are different : no loops in the sequence)

We will say that SA and SB have in common the « ***generalised attribute*** »

V0 – V1 – V2 … VN S

In the previous example, Event 1 and Event 2 have in common the *generalised attribute* :

takes place – is located in – is member of Paris Luxury Districts

This attribute is made with three sentences. The simplest attributes are made with one sentence, like :

lives in Berlin

Given a Semantic Net, we will compute the set of all the generalised attributes which are common to at least two subjects. (This set is finite –no loops).

Now, to each subject, we can associate its generalised attributes.

We can build a matrix SA, such that

SA(i,j) = 1 if subject i has attribute j

SA(i,j) = 0 otherwise.

We have finally transformed our complex semantic networked world into a simple binary matrix.

On this binary matrix, we can now -and with more reasons than on the initial texts- apply "brute force" numerical, continuous processes like statistics and data analysis:

- **compute the distance between two subjects** as a function of their shared generalised attributes. There are in the litterature dizains of proposed distances between two objects sharing boolean properties. For instance some of them take into account the frequency of the attributes: two subjects are closer to each other if they share a scarce atttribute rather than a frequent one.
- **build clusters of subjects,** putting together in the same class subjects sharing enough generalised attributes

This latter process gives strong guidance in the fusion decision process:

- subjects found in the same cluster will be candidates to be merged in an unique one
- subjects found in differents clusters will be candidates to be considered as distinct ones

More subtle decisions will be taken by considering point to point distances between two subjects, and, ultimately, by inspecting the very list of shared and non shared generalised attributes between them.

The chain of processes:

Semantic Net → Generalised Attributes Matrix → Distances and Clusters → Fusion Decision

has the advantage of being very systematic, and to combine two modes:

- an *automatic mode* : compute all attributes, distances and clusters
- a *human mode*: visualise the resulting topology and take fusion decisions

In general, the process will be iterative:

a) a step of initial subjects identification and fusion (as being same or different)
b) computation of the new set of generalised attributes after subjects fusion, yielding a new set of shared attributes
c) new evaluation of distances and clusters
d) new fusion decisions
e) iteration on step a)

The process stops at step d) when no new fusion decisions can be taken.

4.0 IDENTIFICATION OF MORE COMPLEX PHENOMENA

In the previous paragraphs, we addressed the problem of identifying individual subjects: an event, a car, a person.

In the real world, more complex entities exist, from concrete ones to abstract ones :

- groups of people (a terrorist group, a sport club)
- groups of groups (a federation of sport clubs)
- ideologies (the british neo-liberalism)
- phenomenas: "the rise of religious confrontations in the South suburbs of Cairo"

How far is Intelligence Fusion concerned by such concepts ?

It is for instance important to discover that Event 1 is a symptom of Phenomenon A (religious confrontations) and that Event 2 is a symptom of Phenomenon B (political rivalry), event if Event 1 and 2 have many attributes in common.

We will illustrate how to discover complex entities with our approach through an example. We consider a semantic net which contains two categories of subjects:

Persons and *Meetings*

The sentences in the semantic net are of the form :

Person P / present at / meeting M

Meeting M / attended by / Person P

(These subjects may have been identified using the fusion steps explained in the previous paragraphs).

We would like now to discover the possible existence of *groups of people*, of *different kinds of meetings, links between people, between meetings …*

Following our definitions, a group of people is a set of subjects of the Persons category which share a significant set of generalised attributes. Let us look at the possible forms of the generalised attributes:

a) present at Meeting M
b) attended by Person P
c) present at -- attended by Person P
d) attended by -- present at Meeting M
e) present at -- attended by --present at Meeting M
f) attended by -- present at -- attended by Person P
g) present at -- attended by -- present at -- attended by Person P

etc …

Attribute c) means for instance that two persons have in common to attend different meetings, but where the same person P is present.

We call this way of transforming and procession a Semantic Net *"Litteratus Calculus",* to suggest that , in parallel with *scientific calculus* on technical data, a lot of useful computations can be made at a fine grain on data from textual origin.

Imagine now the following situation:

A small political organisation wants to infiltrate large meetings. This organisation is made of cells of a limited number of agents:

Cell A with agents Agent A1, A2, A3
Cell B with agents B1, B2, B3, B4
Each cell responds to a leader (Leader A, Leader B), not member of the cell.
Agents of the same cell are in general present at the same meetings.
There are also Cell Meetings with their Agents and Leaders
A meeting is attended by many Persons.

We have made experiments with Ideliance with simulated data to describe such a situation.

The *only input* of the fusion process is a Semantic Net of sentences like:

Person X / present at / Meeting Y

and indeed no prior knowledge about the structures of groups or roles of persons is available.

The computation of clusters of persons will give the expected results:

Cell A and Cell B will be identified as clusters containing the agents, because, among others and for example :

Agents B1,B2,B3 share the attribute of type c) :

present at -- attended by Agent B4

Agents B1,B2,B4 share the attribute of type c) :

present at -- attended by Agent B3

etc …

We see that a *mesh of relations* links the members of a Cell, and identify it as an interesting result.

Once a cluster is found, the system exhibits the attributes which are shared by most of its members. Thus the result is quite expressive:

Persons B1, …., B4 form a cluster which has in common to be present at meetings with the other members of the cluster.

(theses results are obtained through fixing some threshold to determine how tolerant we want to be on the homogeneity of the groups, and the algorithm works on non perfect situations: not all members of a Cell need to be present at all their meetings …)

Finally, we have discovered several concepts:

- existence of groups of persons, an interesting seed to discover the structure of organisations
- notion of roles of persons
- we can know « which meetings are infiltrated by which cell »

Similarly, we will discover the existence of Cell meetings, and of Cell Leaders: all members of a cell will have in common to be present in meetings with their leader.

Futher steps of analysis could be found:

If we examine ordinary attendees, sympatisers of the infiltrators will be found in clusters which share attributes of type g):

present at -- attended by --present at -- attended by leader A

present at -- attended by --present at -- attended by leader B

In other words: sympatisers of a given organisation often attend meetings where the -masked- infiltrators reporting to the leaders of this organisation are present.

Leaders A and B will themselves be found in the same cluster, characterised by shared attributes of the form

present at -- attended by --present at -- attended by Sympatiser X

This "Leaders cluster" represents, embodies the concept of their organisation.

5.0 CONCLUSION

Our first experiments with Ideliance tend to prove that Semantic Nets can play an important role in Symbolic Intelligence Fusion.

Transforming and merging heterogeneous informations from various formats (databases, tables, messages, texts) into an unique format (what we called *Format Fusion*) is a good basis for *Intelligence Fusion*.

First, it offers an efficient support for "manual" seamless inspection and navigation of the whole set of informations.

Second, it becomes a material upon which powerful data analysis (distances and clusters computation) can be performed, once the *Generalised Attributes* we introduced are computed.

We call this process *"Litteratus Calculus"*.

Our conjecture is that the objects resulting from this analysis form the backbone of the Intelligence Fusion process, which, ultimately is the domain of human decision.

THE CASE FOR USING SEMANTIC NETS AS A CONVERGENCE FORMAT FOR INFORMATION FUSION

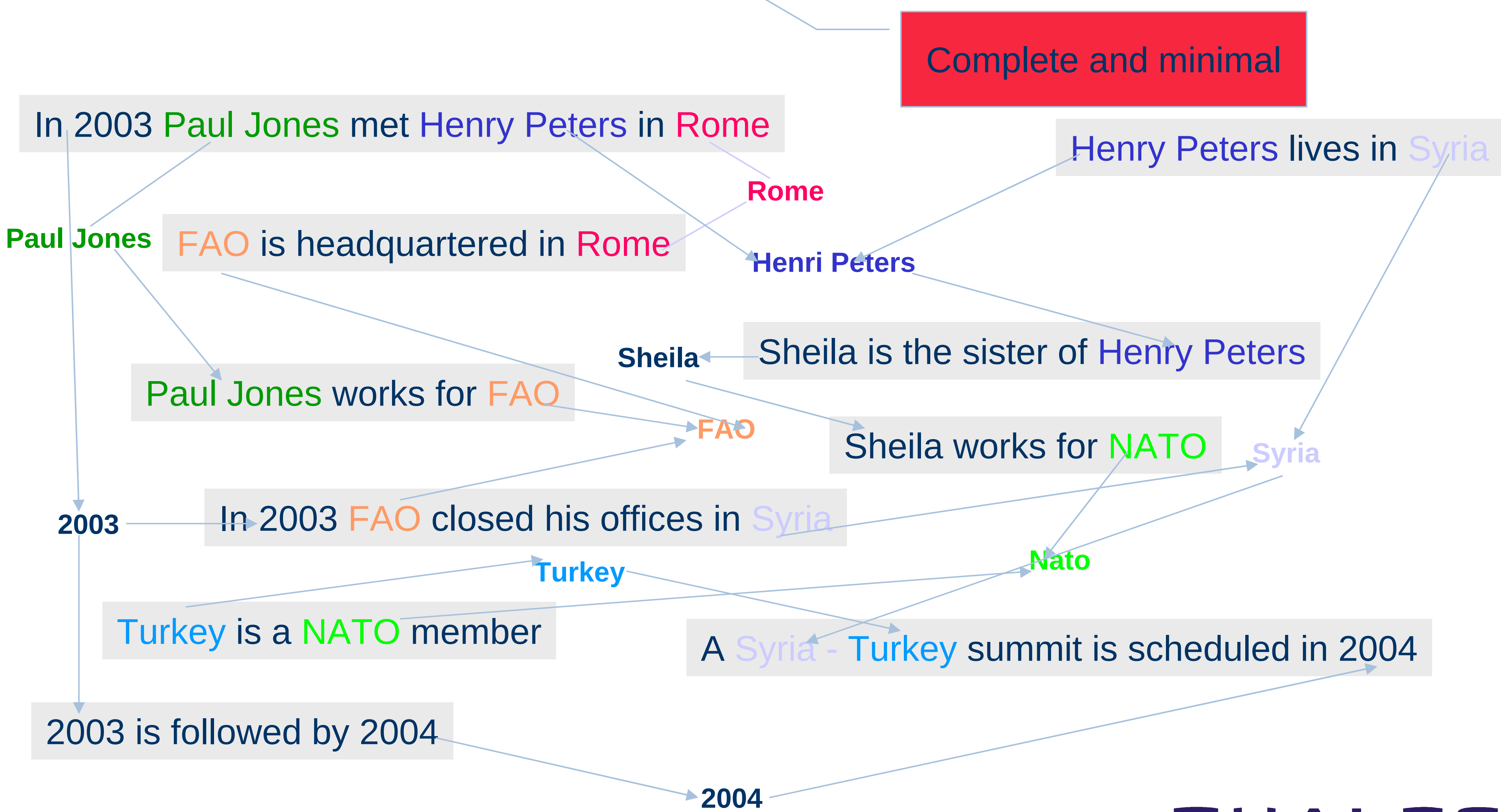
R.T.O. INFORMATION SYSTEMS TECHNOLOGY PANEL SYMPOSIUM ON MILITARY DATA AND INFORMATION

Prague October 21st 2003
Jean Rohmer Thales Communications France

THALES

WHAT ARE SEMANTIC NETS ?

« A SEMANTIC NET IS SIMPLY A SET OF SHORT SENTENCES SHARING WORDS »



WHAT IS UNIQUE WITH SEMANTIC NETS

BOTH HUMANS AND MACHINES UNDERSTAND SEMANTIC NETS

Close enouth to natural language for human understanding

Structured and regular enouth for machine processing

REPRESENT KNOWLEDGE THE SAME WAY

IT IS REPRESENTED IN OUR **BRAINS**

no texts, no databases, but small pieces of interrelated knowledge

MACHINES CANNOT UNDERSTAND USUAL LANGUAGE

USUAL PEOPLE CANNOT UNDERSTAND DATABASE PROGRAMS

GOOD NEWS ABOUT SEMANTIC NETS

SEMANTICS IS THE N°1 « STANDARD » ADOPTED BY MANKIND

==> SEMANTIC-BASED INFORMATION SYSTEMS

WILL **OUTPERFORM** CLASSICAL ONES

IN TERMS OF INTEROPERABILITY

SHARE PART OF OUR KNOWLEDGE WITH MACHINES

PERFORM POWERFUL COMPUTATIONS ON SEMANTIC NETS

AMPLIFY OUR INTELLIGENCE

CHEAPER , **FASTER**, **EASIER**

BECAUSE **USERS CAN TAKE PART IN THEIR CONSTRUCTION**

THALES

SEMANTIC NETS ARE NOT « FEATURES » OF A SYSTEM

They can represent Data, Informations, Knowledge, Ontologies, Rules, Templates, Behavioural Models, Theories …

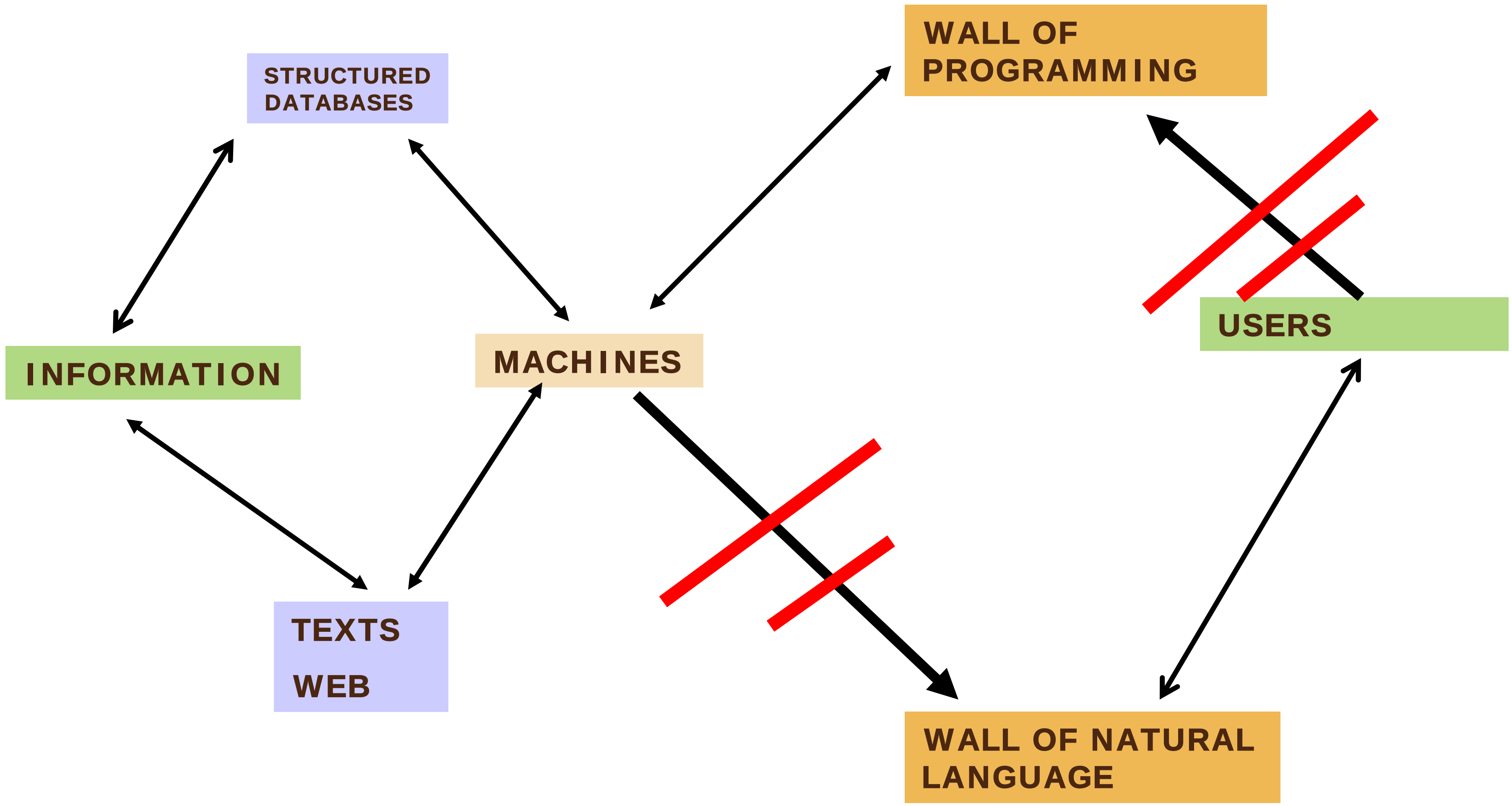
They can be used to perform DS, Inference, Clustering, Discovery, Analogy, Rhetorics …

Our Brain is a champion at Information Fusion

Automatic systems for fusion should mimic our brain,

starting with its means of representation

« CLASSICAL I.T. » IS AN OBSTACLE TO FUSION



THALES

LESSONS LEARNT FROM YESTERDAY SESSIONS

JOINT DATABASE MODEL

DESIGN AND SHARING IS IMPOSSIBLE

DECISION:

DO NOT SHARE MODELS

DO NOT USE MODELS

APPLICATIONS PROGRAMMING BY COMPUTER SCIENTISTS IS

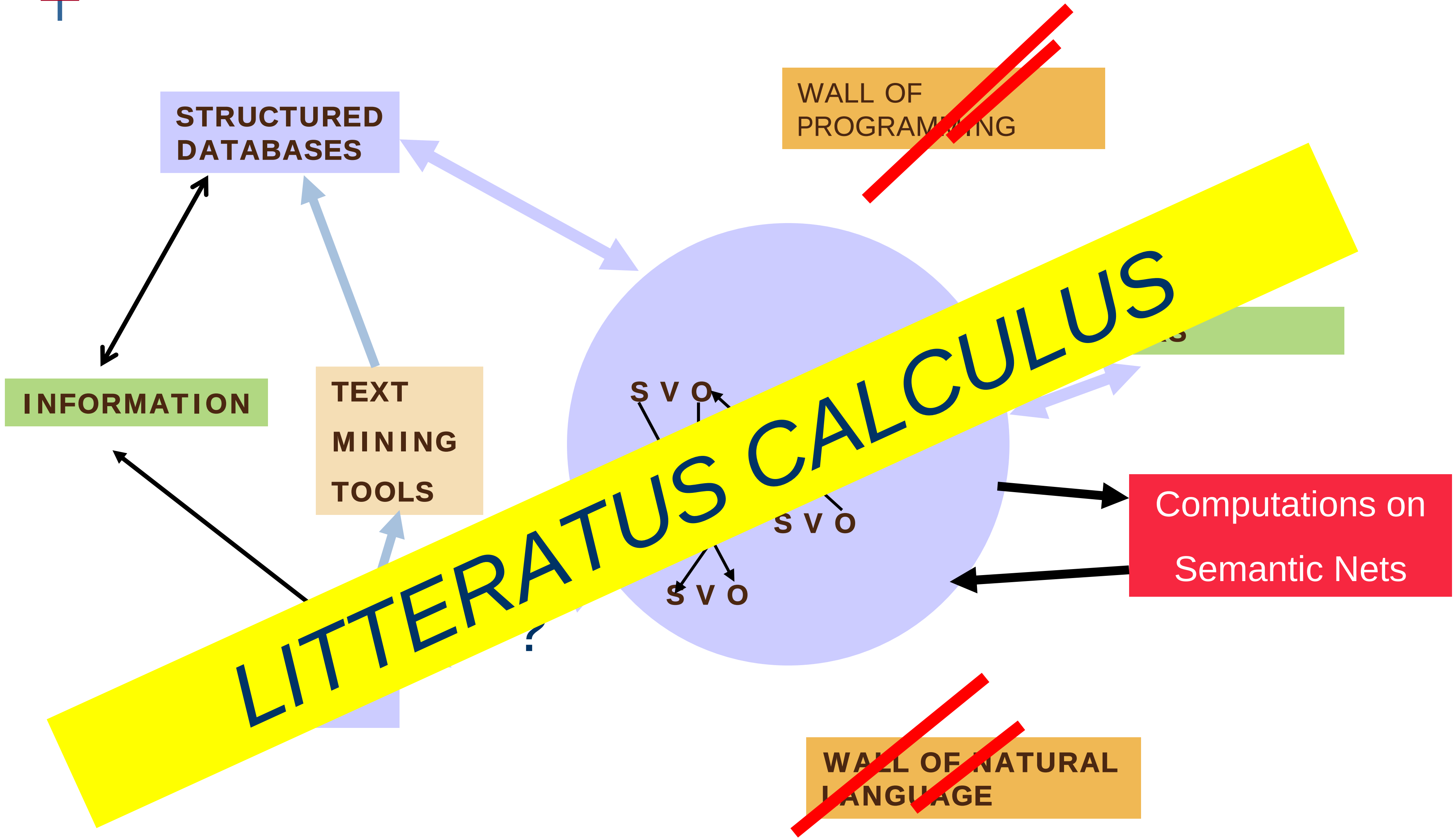
A COSTFUL NIGHTMARE

DECISION:

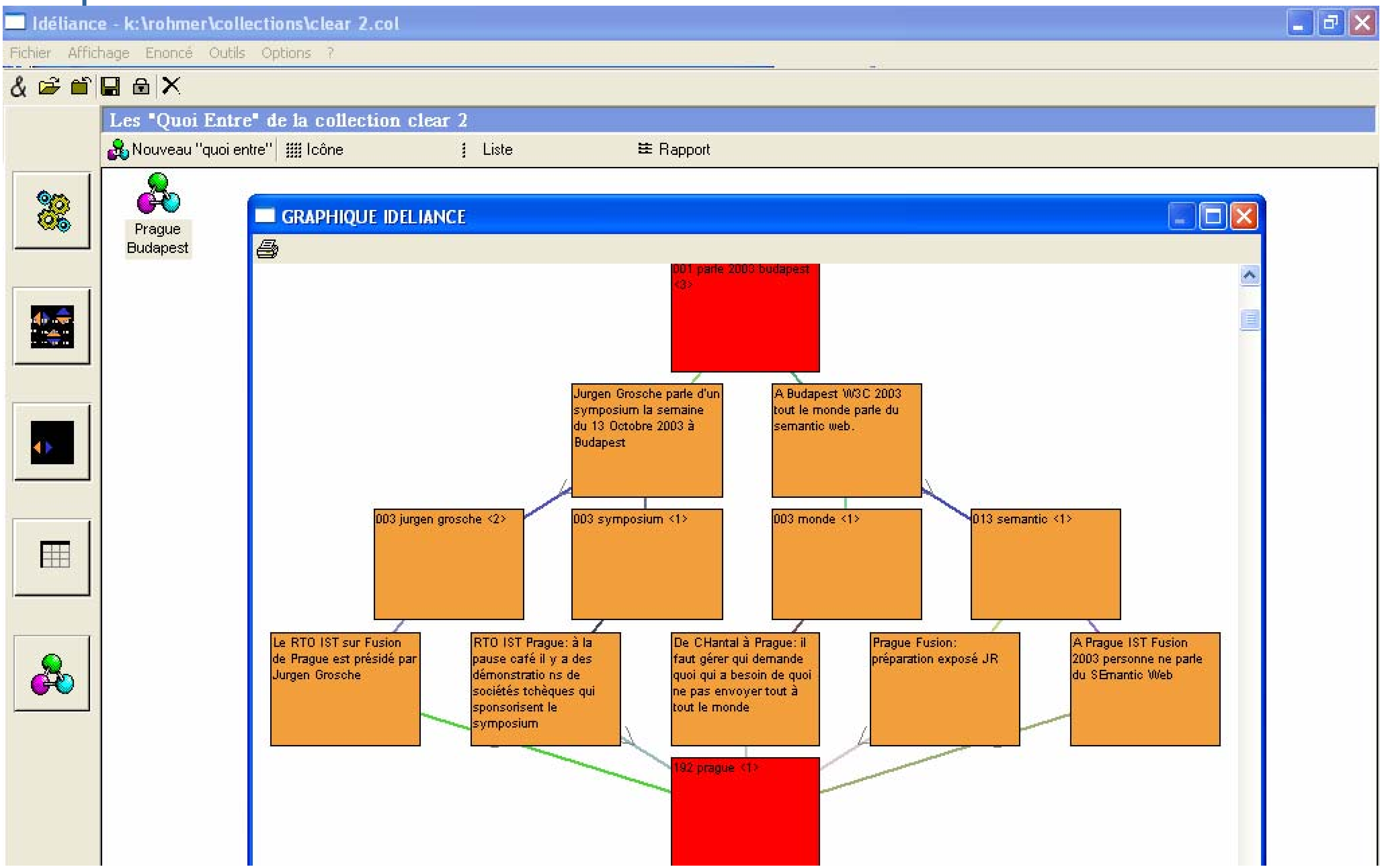
DO NOT PROGRAM

DO NOT USE PROGRAMMERS

FORMAT FUSION WITH SEMANTIC NETS



PRAGUE, BUDAPEST AND ME



FORMAT FUSION

VS

INTELLIGENCE FUSION

La capacité de recevoir des représentations des objets par la manière dont ils nous affectent s'appelle la ***sensibilité***. C'est au moyen de la sensibilité que les objets nous sont donnés, et elle seule nous fournit des intuitions

Mais c'est par l'***entendement*** qu'ils sont ***pensés,*** et c'est de lui que sortent les ***concepts***

Toute ***pensée*** doit aboutir en dernière analyse, soit directement, soit indirectement, à la ***sensibilité*** qui est en nous, puisque aucun objet ne peut nous être donné autrement.

Notre connaissance dérive de deux sources, la capacité de *__recevoir des représentations__* et la faculté de *__connaître__* cet objet au moyen de ces représentations.

Par la première un objet nous est *__donné__*, par la seconde, il est *__pensé__* dans son rapport à cette représentation

Intuition sensible et entendement:

Ces deux capacités ne sauraient échanger leurs fonctions: l'entendement ne peut rien percevoir ni les sens rien penser.

La Connaissance ne peut résulter que de leur union.

Aussi distinguons-nous la science des règles de la sensibilité en général, ou ***esthétique,*** de la science des règles de l'entendement en général, ou ***logique***

L’exactitude et la précision des connaissances sont plutôt funestes en général

Il est rare en effet qu’elles remplissent d’une manière adéquate la condition de la règle.

En outre elles affaiblissent ordinairement cette tension de l’entendement nécessaire pour apercevoir les règles dans toute leur généralité et indépendamment des circonstances particulières

EMMANUEL KANT 1781

Example of Intelligence Fusion : Networks Mapping
(of people, of communications)

Problem:

Discover the structure of an organisation -human or technical- from sparse indices

Solution

A) Merge various Information sources into a simple semantic network

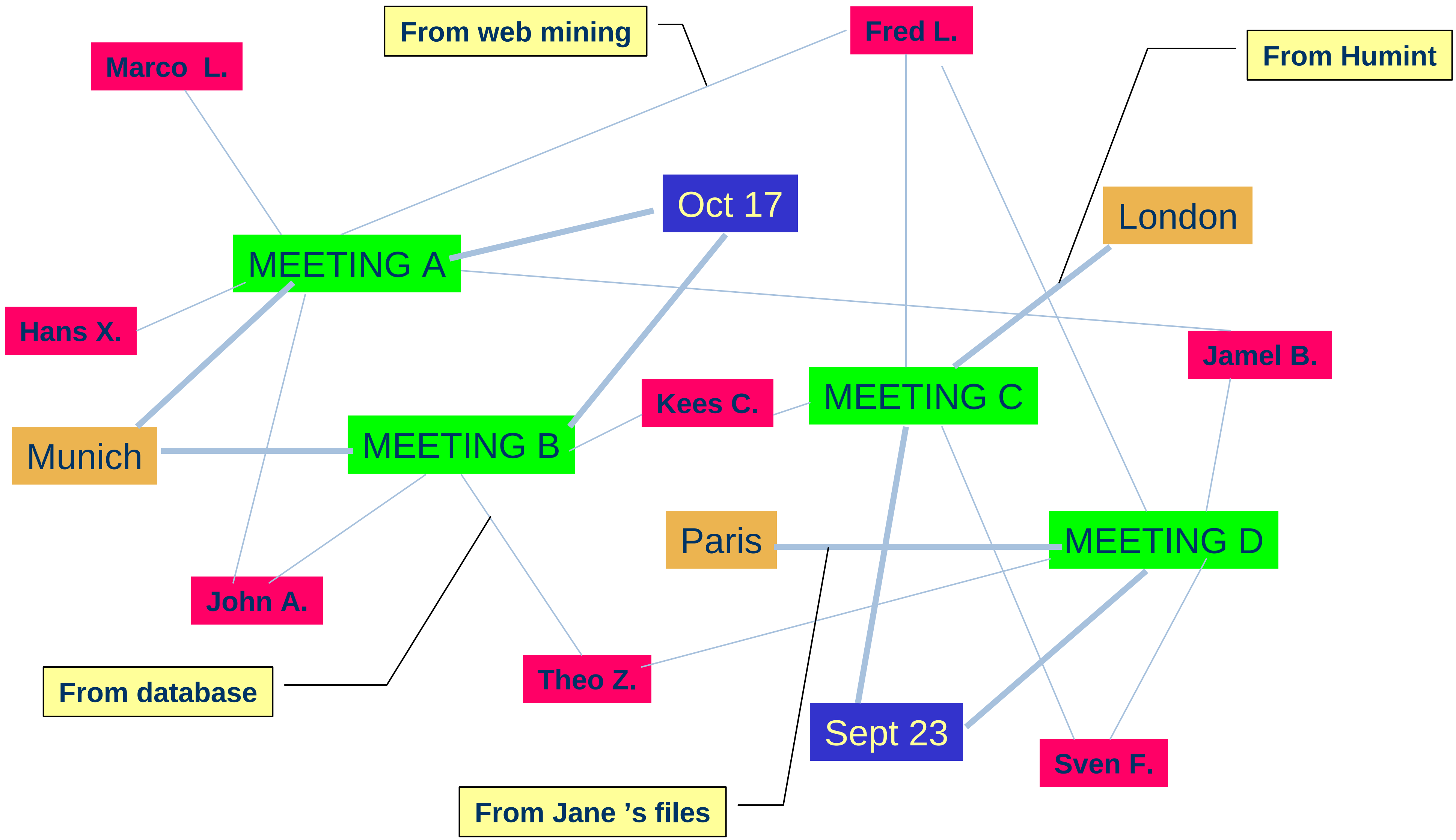
B) Perform clustering (of persons, of places, of meetings)

to discover homogeneous classes of behaviours

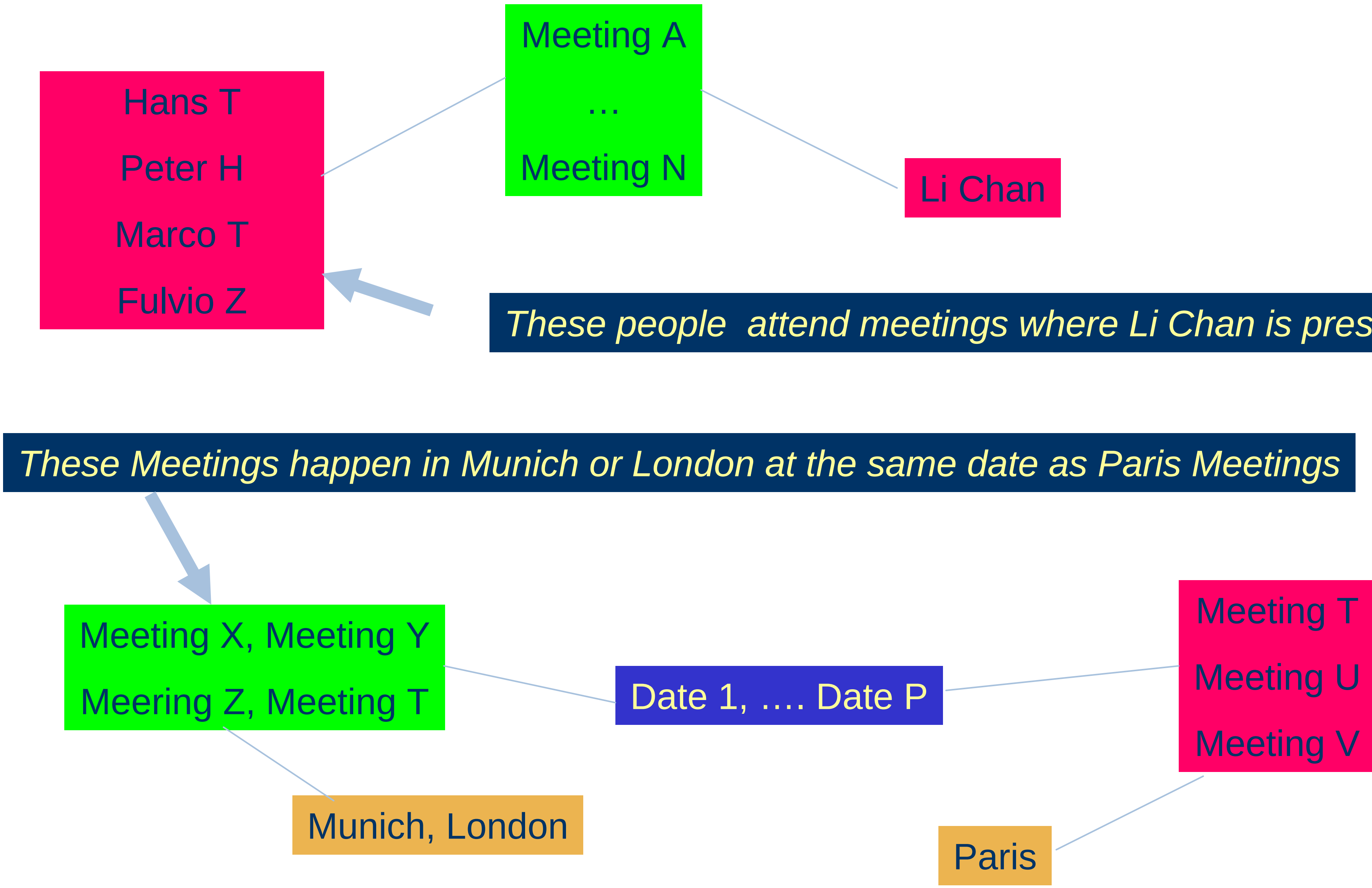
C) Interpret the resulting classes in term of organisations

by using expert rules

STEP A: MERGE MANY SOURCES INTO AN UNIQUE SEMANTIC NETWORK OF MEETINGS, PEOPLE, PLACES, DATES ...



STEP B: AUTOMATIC DISCOVERY OF CLUSTERS OF *SUBGRAPHS* SHARING PROPERTIES



STEP C: APPLY KNOWLEDGE TO UNDERSTAND ORGANISATIONS AND THEIR BEHAVIOURS

EXPERT RULE 1:

IF a person regularly attends meetings in various towns on the same day

THEN this person is a big boss

EXPERT RULE 2:

IF more than « 6 » persons regularly meet with another one

THEN the first ones belong to a **«** cell **»**

AND the second one is the leader of the cell

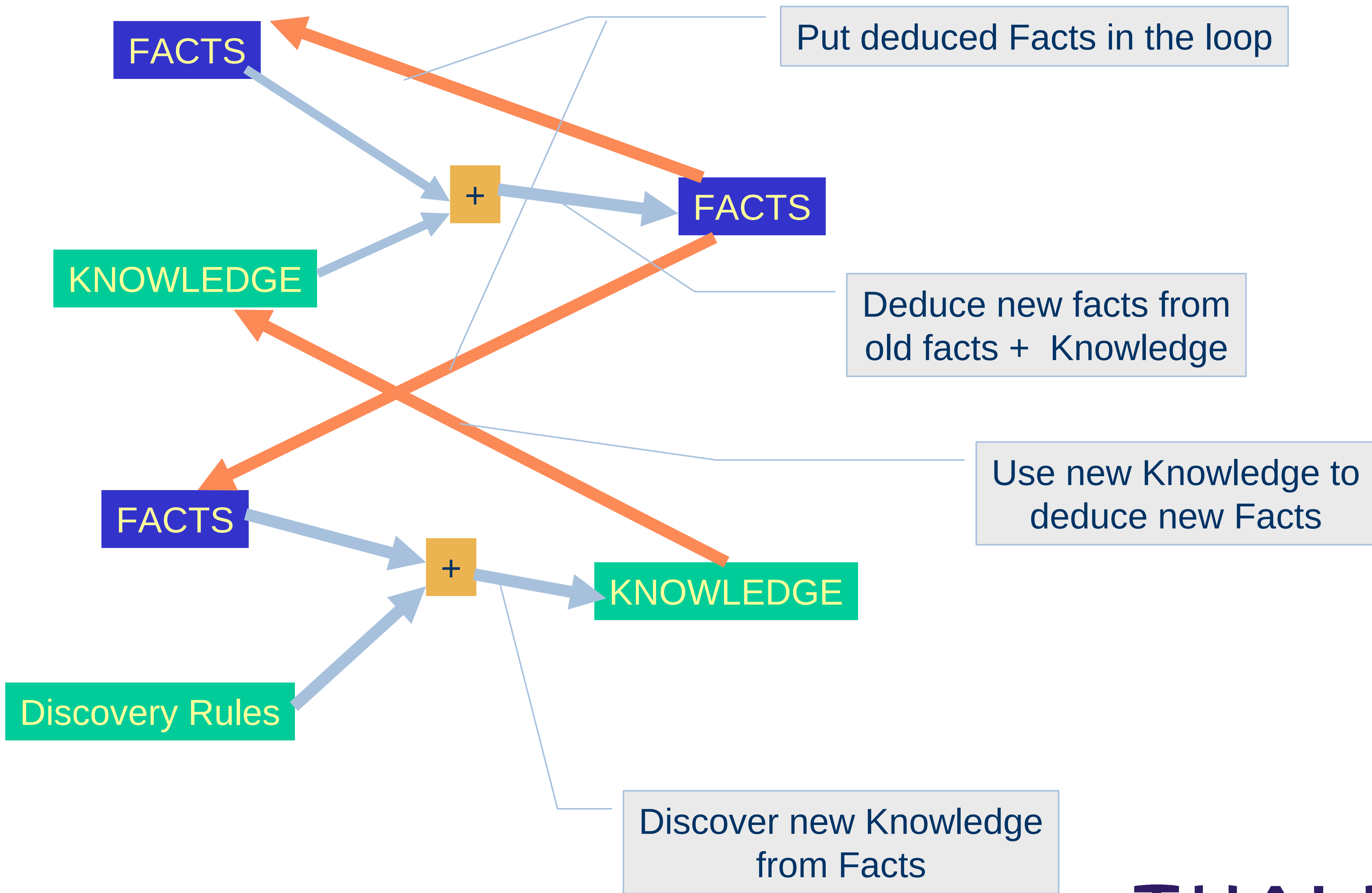
EXPERT RULE 3:

IF members of cell C regularly attend meetings organised by Organisation A and Cell C belongs to Organisation B

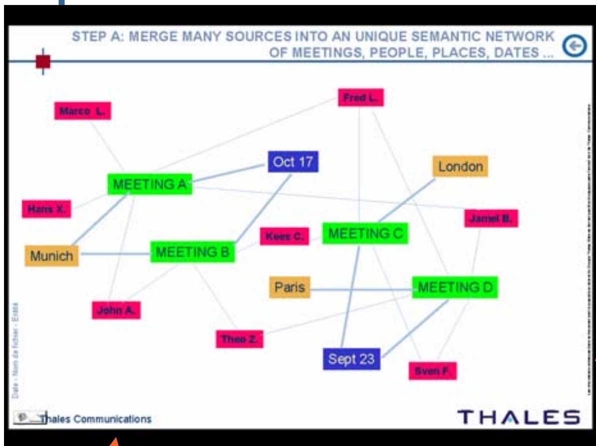
THEN B is trying to infiltrate A

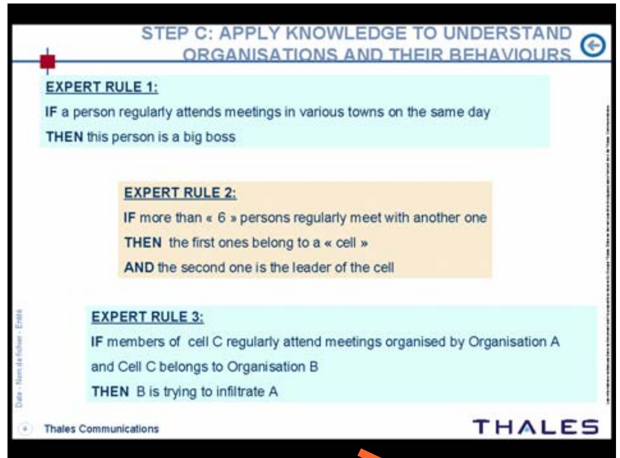
THALES

THE « FACTS - KNOWLEDGE - DISCOVERY » LOOP

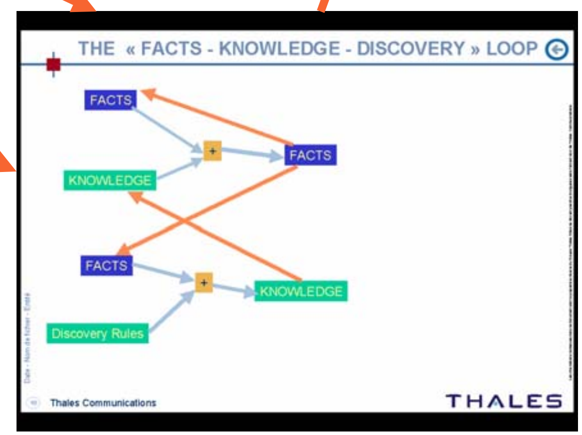


Summary: Mapping of an Organisation via a Journey from Informations to Intelligence through Semantic Representation





Intelligence
Report
About
Organisation XXX

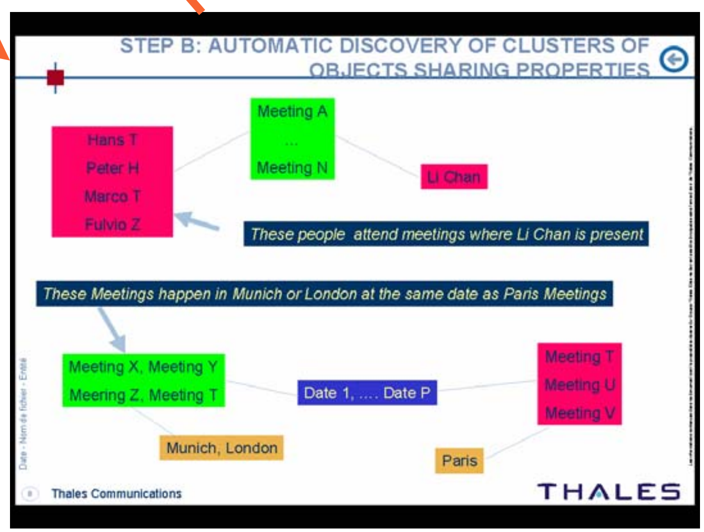


Raw

Informations

About

Organisation XXX



THALES

USE *SEMANTIC NETS* AS BASIC

–*not* optionnal-

REPRESENTATION FRAMEWORK

AND PUT THEM AT WORK !!!!

USE *SIMILARITY* AS BASIC

-not optionnal-

CALCULUS FRAMEWORK

THERE IS NO SUBSTITUTE TO HARD WORKING

Thomas Edison

LE CHEMIN N'EST PAS DIFFICILE

MAIS « *DIFFICILE* » *EST* LE CHEMIN

Michel de Montaigne